Semi-Automatic Creation of an Emotion Dictionary Using WordNet and its Evaluation

David B. Bracewell
GE Global Research
Niskayuna, NY USA

***Abstract*—This paper describes an algorithm for semi-automatically creating an emotion dictionary using WordNet and evaluating the created dictionary. The algorithm takes as input a set of seed words that have been manually assigned with WordNet senses and emotion information. An initial dictionary is automatically created using the seed words and WordNet. Then, various correction stages are performed where parts of the dictionary are shown to the user for verification. An almost 6,000 word sense emotion dictionary was created from only 130 seed words using the proposed algorithm. Evaluation of the created dictionary shows that it has a good vocabulary coverage for an independently tagged emotion corpus. The created dictionary was also used to classify the opinion of news articles. A simple algorithm that relied completly on the strength of the created dictionary achieved an accuracy of 84%.**

***Index Terms*—Natural Language Processing, Affective Computing, Emotions, Dictionary Construction**

I. Introduction

Emotion plays an important role in everyday life. Currently, standard language processing ignores the emotional message that is interlaced in much of human dialog and instead, only focus on the explicit message. However, the emotion behind the dialog plays an equally important role in understanding the true or implied meaning intention of the speaker/writer.

In order to answer the call of understanding emotion, *affective computing* was born. Recently, affective computing research has joined with natural language processing research to tackle text based affect analysis, such as [1], [2] and [3]. However, this joint venture is still in its infancy. As such, resources that are necessary for proper analysis, such as dictionaries, corpora, etc., are scarce.

To deal with the scarcity of resources, this paper describes a semi-automatic approach to creating emotion dictionaries using WordNet. WordNet is a lexical dictionary that groups words with synonymous meanings into groups and treats them as concepts [4]. It has quickly become a standard tool used in Natural Language Processing, Information Retrieval and other fields that need the lexical knowledge it contains. The algorithm to create emotion dictionarie combines the cognitive knowledge in WordNet with a set of user given seed words to create an expanded emotion dictionary. Using this approach, a close to 6,000 word sense emotion dictionary has been created. The validity of the dictionary will be shown by evaluating the coverage of the dictionary and the usability of the dictionary in a real world task: opinion classification.

This paper will continue as follows. First, section II will give background and related work. Next, in section III the basic emotion categories used are described. Then, section IV defines the structure of the dictionary including the information it contains. After that, section V will give the details of the emotion dictionary creation algorithm. Next, section VI will give an evaluation of the created dictionary. Finally, section VII will discuss future work and give concluding remarks.

II. Related Work

Research on affect analysis has received increased focus in the past few years. There has been much research done, but most of it on classification and applications. Research on creation of affective resources (i.e. dictionaries, corpora, etc.) is sparse and, as such, the resources available are still limited. The following paragraphs will take a look at some of the research done on resource creation and applications using affective analysis.

One approach to creating affective vocabularies is to use human participants, such as Cowie et al. [5]. While these vocabularies are grounded on human perception of opinion they are costly to create and update. This may not be practical for many researchers. Updating the vocabulary with new words or emotions would require another set of participants, which could possibly have a different emotional background. Moreover, local cultural influences can creep in if there is not a diverse enough pool of participants.

A related field that has seen a great deal of work in resource creation is sentiment, or opinion, analysis, where the goal is to determine if text carries a "+" or "-" (which is referred to as polarity) feeling. Esuli and Sebastiani assigned positive, negative, and objective scores to all of the senses in WordNet [6]. Kanayama and Nasukawa created an unsupervised approach using context coherency for extracting domain-dependent polar terms from unannotated corpora that achieved 94% precision [7]. However, for affect analysis, polarity is typically not enough, which has lead to researchers using manually created lexicons.

Applications of affect analysis include affective retrieval, instant messaging, etc. Subasic and Huettner examined fuzzy semantic typing for affect analysis in retrieval systems [3]. The lexicon they used was created, as they say, in a "haphazard" manner; one of their goals in the future was to semi-automatically create the lexicon so that it would be more robust. Ma et al. looked at estimating the emotion of text for Internet based chat applications [1]. Their algorithm used a keyword spotting technique and required system designers to

TABLE I
THE 14 BASIC EMOTION CATEGORIES

Basic Emotions	Sub-Emotions				
Anger	Rage	Envy	Torment	Irritation	Hate
Disgust	Disrespect				
Sadness	Suffering	Disappointment	Neglect	Pity	Sympathy
Shame	Regret				
Fear	Horror	Anxiety			
Apathy					
Surprise	Confusion				
Anticipation					
Calmness					
Joy	Contentment	Optimism	Relief	Enthrallment	Cheerfulness
Love	Affection	Lust			
Desire	Longing				
Courage					
Praise	Congratulate	Respect	Approve	Gratitude	

adjust the emotion database manually.

III. BASIC EMOTION CATEGORIES

This first step in building an emotion dictionary is deciding on the basic emotion categories. There is no one correct answer to what the basic emotion categories should be. Everyone from philosophers to psychologists have stated their opinion, but there is still no agreed upon standard. As such, often times, the categories vary between research to suit the needs of the application.

While the dictionary creation algorithm presented in this paper is not tied to any one emotion theory, a set of basic categories have been defined to suit our needs. The chosen categories draw heavily from the ideas of Parrot [8], who classified emotion into a shallow tree structure.

Table 1 shows the 14 basic categories and some of their subcategories. The first 6 basic emotions have a polarity of "-" and the last 6 have a "+" polarity. Surprise and Anticipation have a specially defined value of "?" denoting that the category itself carries no polarity. The words in the "?" category have their polarity decided on an individual basis and possibly change in different context.

IV. EMOTION DICTIONARY STRUCTURE

The dictionary is made up of information that is useful for both affect and sentiment analysis. The information can be used for both shallow surface level approaches or more involved natural language processing based approaches. Each entry in the dictionary contains the word, part-of-speech, WordNet sense id, emotion category, polarity, emotion probability information and WordNet gloss, which is a dictionary definition for the word. An example of the dictionary structure can be seen in figure 1.

V. EMOTION DICTIONARY CREATION ALGORITHM

The emotion dictionary creation algorithm is broken up into five steps: Initial Creation, Multiple Category Correction, Category Outlier Correction, Polarity Outlier Correction and Probability Calculation. The algorithm requires as input a set of seed words with their WordNet sense and an assigned emotion category. The following subsections will take a look at each of the parts in more detail.

A. Initial Creation

The initial creation of the emotion dictionary is automatically done using the given seed words. In this paper, 130 emotion words were extracted from Parrot's [8] classification; emotion category and WordNet sense(s) were assigned manually and were fed into the initial creation step. An overview of the initial creation step can be seen in figure 2.

The seed words are stored in a queue and processing continues until the queue is empty. Each word taken from the queue, as a seed word, is looked up in WordNet using the word and sense information. The synset, hyponyms and derivationally related words are extracted and have the emotion category of the seed word assigned to them. Word senses that have been previously seen and assigned with the current emotion are discarded to prevent duplicate entries in the dictionary.

The words in the synset and hyopnym set are then enqueued to become new seed words. This process allows for extraction of the full hyponym tree as well as any related forms, i.e. all parts of speech are possible. Using the 130 seed words, a little over 6,100 word senses were extracted and added to the initially created emotion dictionary.

B. Multiple Category Correction

The next step is to look at words assigned to multiple emotion categories for possible miscategorizations. Certain words should have multiple emotion categories assigned. For example, sense 1 of the noun "apprehension", which WordNet defines as "fearful expectation or anticipation" [1] should be assigned the emotion categories "fear" and "anticipation."

[1]http://wordnet.princeton.edu/perl/webwn?s=apprehension

Word: love	**Part-of-Speech:** n	**Sense:** 1	**Emotion:** love	**Polarity:** +	**P(E):** 0.800	**P(This E):** 0.750
Gloss: a strong positive emotion of regard and affection; “his love for his work”; “children need a lot of love”						
Word: low_spirits	**Part-of-Speech:** n	**Sense:** 1	**Emotion:** sadness	**Polarity:** -	**P(E):** 1.000	**P(This E):** 1.000
Gloss: a state of mild depression						
Word: lucky	**Part-of-Speech:** a	**Sense:** 2	**Emotion:** joy	**Polarity:** +	**P(E):** 1.000	**P(This E):** 1.000
Gloss: having or bringing good fortune; “my lucky day”						
Word: ruffle	**Part-of-Speech:** a	**Sense:** 3	**Emotion:** anger	**Polarity:** -	**P(E):** 0.167	**P(This E):** 0.500
Gloss: disturb the smoothness of; “ruffle the surface of the water”						

Fig. 1. Example Dictionary Structure

Fig. 2. Initial Emotion Dictionary Creation Algorithm Overview

However, having multiple emotion categories should not be typical for most words.

The words assigned with multiple basic emotions are automatically extracted from the initially created dictionary. They are displayed to the user along with their assigned emotions and their WordNet gloss. The user then manually decides what to do for each of the word senses. The dictionary constructed in this paper had just under 300 word senses with multiple emotion categories assigned. Of these 300 senses, 40 were left as is and the remaining were modified.

C. Category Outlier Correction

It is a strong possibility that the differences in motivations behind the creation of WordNet and the emotion dictionary will cause the WordNet hierarchy to introduce errors when assigning emotion categories automatically. Because of this, the category outlier correction step is performed. It tries to determine possible miscategorized words senses.

This is done by first performing k-means clustering on the dictionary. To perform the clustering, each word is converted into a vector representation where the dimensions of the vector are the basic emotion categories. The values are determined by the number of word senses in the synset, hypernym, hyponyms, and derivationally related forms that have that emotion. K-means is then performed 100 different times with k equal to the number of emotion categories using Euclidean distance as the distance metric.

After completion of each run of the k-means algorithm the most predominate emotion in each cluster is assigned as the cluster's emotion. Using the 100 runs, the most probable emotion category for each word sense can be estimated by choosing the category that the word sense was most frequently clustered into, e.g. if the word sense was a part of a "joy" cluster 80 times and a "love" cluster 20 times then its most probable emotion category is "joy."

Those word senses whose predicted emotion category differ from their initially assigned category are flagged for manual verification. The user then makes the choice if the emotion category should be changed or not. For the dictionary created in this paper, there were 98 possible miscategorized word senses found and 11 were actually changed.

D. Polarity Outlier Correction

The final step is polar outlier correction. The initial polarities are assigned based on the emotion category of the word sense. In a similar fashion to category outlier correction, k-means clustering is used to determine possible problems in the assigned polarity.

Each word sense is converted into a three dimensional vector ('+", "-" and "?" dimensions). The values for the dimensions are determined by the polarity of the word senses in the synset, hypernym, hyponyms and derivationally related forms. 100 runs of the k-means clustering algorithm, with k equal to three and using Euclidean distance as the distance metric, are performed. Similarly to Category Outlier Correction, the most probable emotion category for each word sense can be estimated by choosing the category that the word sense was most frequently clustered into over the 100 runs.

The polarity of each word sense is predicted and compared to the initially assigned polarity with non-matches being given to the user for manual verification. For the dictionary created in this paper, this resulted in 20 possible errors, 9 of which were actual errors. All 9 of these were initially assigned a "?" polarity.

E. Emotion Probability

Each emotion carrying word in the dictionary was assigned two different probabilities. The first was the probability that the word carries emotion. This can be calculated as the number of word senses that carry emotion divided by the total number of senses for the word, shown in equation 1. In this case each sense is treated as equally probable, which may or may not be the case in real use. However, it should give a fair estimate.

$$P\left(E|W\right) = \frac{\textit{Count(}s_i \in W \textit{ that carry emotion)}}{|W|} \quad (1)$$

The second probability calculated is the probability that the word has the given emotion. Some words can carry multiple emotions and this probability helps to determine which emotion is the most probable. It is calculated by simply dividing the number senses given the emotion divided by the total number of senses carrying any emotion. This calculation can be seen in equation 2.

$$P\left(e_{given}|W\right) = \frac{\textit{Count(}s_i \in W \textit{ that carry } e_{given}\textit{)}}{\textit{Count(}s_i \in W \textit{ that carry emotion)}} \quad (2)$$

VI. Evaluation

The emotion dictionary created using the algorithm detailed in this paper is made up of 5,956 word senses covering all 14 basic emotion categories. Testing the quality of the dictionary is a difficult task. In order to do so, we look at two evaluations. The first looks at the dictionary coverage using an independently annotated emotion corpus. The second uses the dictionary in analyzing the opinion of news articles.

A. Dictionary Coverage

The dictionary was created from a controlled source and the seed words were manually chosen and known to be emotions. Because of this, the dictionary should have few, if any errors. Therefore, the main question becomes the coverage of the dictionary's vocabulary.

To test the coverage, an emotion corpus was used that was annotated independently of the dictionary creation process. The corpus is made up of 1,000 English sentences from a contrastive dictionary of Japanese-English emotion expressions [9]. The sentences had a total of 1,575 tagged emotion words. Each of the words were looked up in the created dictionary to check for existence. 88% of the words were found in the created emotion dictionary. Of the 12% not found, 3 words were in WordNet but not in the dictionary and the rest were idiomatic phrases not in WordNet.

B. Opinion Classification for News Articles

The next evaluation of the created dictionary used a real word task, opinion classification. In this evaluation we looked at classifying the overall opinion in news articles. This type of research is not new and has most recently been done by [2]. First, a brief overview of how classification was performed will be given and then the details of the evaluation will be shown.

1) Classification Process: The algorithm for classification does not make use of machine learning. Instead, it is solely reliant on the emotion dictionary which allows for the validity of the dictionary to be judged. Each polarity (opinion) is given a score based upon the words carrying that opinion in the text. Only the traditional "+" and "-" polarities are used and the "?" polarity introduced in this paper is ignored.

The first step in classification is to break the articles up into words. Starting with the first word, each word is analyzed in the context of the next four words. This four word window allows emotional phrases of up to five words to be identified in the text. Four was taken as the window size based on analysis of the emotion words/phrases in the created dictionary. Each of the phrases are passed through the morphological analysis routine used in WordNet to normalize the phrase. The longest phrase found in the dictionary is chosen.

If an emotion phrase is found then a window of three words to the left, chosen based on analysis of the distance of negations when used with emotion words/phrases, are examined for possible negations of the emotion. Emotion negations can come through the use of "not," "no," "'nt," etc. The polarity associated with the word (or its opposite, if negated) score is incremented by the probability of that the looked at word/phrase is an emotion. This is continued for each word in the article. After text processing is completed, the polarity with the highest score is chosen as the opinion of the article.

2) Evaluation: 50 news articles were manually assigned an "+" or "-" opinion and used to test the classification. The articles covered all areas of news from sports, politics, entertainment, etc. 42 (84%) of the articles were assigned the correct opinion. This shows that the dictionary can possibly be useful in classification of opinion when used solely by itself.

C. Discussion

Overall, the presented opinion classification system worked well. Only 8 of the 50 articles had the wrong opinion classified assigned. Two of these 8 were articles about the death of someone famous, in which a recount of the person's life and accomplishments were given. This fond look back overpowered the overall sad situation and caused a misclassification that is easily correctable if using machine learning.

The other six misclassified articles were business related. For this domain of articles it maybe necessary to have a domain specific dictionary beyond just an emotion dictionary. The domain specific dictionary would add implied emotions for common domain words. For example, "foreclosure" would have the emotion "sadness" assigned to it.

VII. Conclusion and Future Work

This paper examined an algorithm for semi-automatic creation of an emotion dictionary using WordNet. Dictionaries created by this algorithm are useful not only in emotion classification, but also sentiment analysis. Through evaluation, we found that by using only a small set of seed words we could achieve a good vocabulary coverage for a manually annotated corpus. Additionally, it was shown that by using only the dictionary we could achieve an 84% accuracy in classifying the opinion of news articles. We believe that this approach can help researchers quickly create emotion dictionaries for their research.

In the future, we hope to expand the dictionary by adding more seed words to the system. We also are thinking about augmenting the system with some type of machine translation and manual check in order to create a bilingual or multilingual emotion dictionary. Finally, the method was designed for emotion dictionary creation, but the ideas used in the algorithm are not limited to this. As such, we also hope to look at modifying the method to semi-automatically create domain-dependent dictionaries.

References

[1] C. Ma, H. Prendinger, and M. Ishizuka, "A chat system based on emotion estimation from text and embodied conversational messengers." in *Proceedings of the 4th International Conference on Entertainment Computing*, 2005, pp. 535–538.

[2] D. B. Bracewell, J. Minato, F. Ren, and S. Kuroiwa, "Determining the emotion of news articles," in *Proceedings of the 2006 International Conference on Intelligent Computing (ICIC 2006)*, ser. Lecture Notes in Artificial Intelligence, vol. 4114, 2006, pp. 918–923.

[3] P. Subasic and A. Huettner, "Affect analysis of text using fuzzy semantic typing," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 4, pp. 483–496, August 2001.

[4] G. Miller, "Nouns in wordnet: A lexical inheritance system," Princeton, Tech. Rep., 1993.

[5] R. Cowie, B. A. E. Douglas-Cowie and, J. Taylor, A. Romano, and W. Fellenz, *What a neural net needs to know about emotion words*. World Scientific & Engineering Society Press, 1999.

[6] A. Esuli and F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining," in *Proceedings of the fifth international conference on Language Resources and Evaluation*, 2006, pp. 22–28.

[7] H. Kanayama and T. Nasukawa, "Fully automatic lexicon expansion for domain-oriented sentiment analysis," in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP2006)*, 2006, pp. 355–363.

[8] W. Parrot, *Emotions in Social Psychology*. Psychology Press, 2001.

[9] I. Hiejima, *Nichi-Ei Taishoo Kanjoo Hyoogen Jiten*. Tokyodo Shuppan, 1995.